## Classification Network mlp (64,64) mlp (64,128,1024) input feature max mlp input points transform transform pool (512,256,k) 1024 nx64 nx64 nx3 nx1024 shared shared global feature output scores point features output scores 64x64 transform nx128 nxn n x 1088 shared shared matrix multiply matrix mlp (512,256,128) mlp (128,m) Segmentation Network

图 1: PointNet

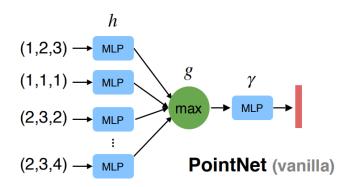


图 2: PointNet

• 排列不变性: 对称函数

$$f(x_1, x_2, \dots, x_n) = \gamma \circ g(h(x_1), \dots, h(x_n))$$
$$\left| f(S) - \gamma \left( \max_{x_i \in S} \{h(x_i)\} \right) \right| < \epsilon$$

可以任意逼近所有连续的集合函数(没找到证明材料)。

- 旋转不变性: Spatial Transformer Network, 并在 loss\_function 上加上这个约束  $L_{reg} = \|I AA^T\|_F^2$
- 点与点之间存在影响: 将全局特征与每个点特征叠加

**分类任务** 使用 T-Net 网络对输入特征进行旋转,然后使用 MLP 使每个点的特征增加,随后对特征进行旋转。这里 PointNet 用的 MLP 是卷积核,将每个点都进行特征提取,使其 channel 变多,本质是信息的冗余。随后使用 Maxpool, 提取整体特征。最后使用 MLP(Linear), 进行分类。

**语义分割** 在分类任务提取整体特征的基础上,将整体特征重复 n 次,然后和每个点的特征叠加起来,最后还是一连串的 MLP(卷积)+softmax